



# The Truth About Better Understanding?

Lewis Ross<sup>1</sup>

Received: 29 December 2019 / Accepted: 8 February 2021 / Published online: 15 April 2021  
© The Author(s) 2021

## Abstract

The notion of understanding occupies an increasingly prominent place in contemporary epistemology, philosophy of science, and moral theory. A central and ongoing debate about the *nature* of understanding is how it relates to the truth. In a series of influential contributions, Catherine Elgin has used a variety of familiar motivations for antirealism in philosophy of science to defend a non-factive theory of understanding. Key to her position are: (1) the fact that false theories can contribute to the upwards trajectory of scientific understanding, and (2) the essential role of inaccurate idealisations in scientific research. Using Elgin’s arguments as a foil, I show that a strictly factive theory of understanding has resources with which to offer a unified response to *both* the problem of idealisations and the role of false theories in the upwards trajectory of scientific understanding. Hence, strictly factive theories of understanding are viable notwithstanding these forceful criticisms.

## 1 Introduction

The notion of understanding has come to occupy an increasingly prominent place in contemporary epistemology, philosophy of science, and moral theory. A central and ongoing debate about the *nature* of understanding is how it relates to the truth. While truth—or knowledge, which entails it—has long been thought to be the primary epistemic goal, some theorists now think a proper appreciation of the importance of understanding calls for a less exacting preoccupation with factivity.<sup>1</sup>

The recent work of Catherine Elgin exemplifies the idea that we should ‘dethrone’ the centrality of truth when theorising about understanding and the goals

<sup>1</sup> This paper focuses on the relationship between understanding and truth. For recent defences of the idea that understanding necessarily requires *knowledge* see, for example, Sliwa (2015, 2017) and Kelp (2015, 2017), with Grimm (2006) and Ross (2020) providing helpful discussion. Another debate concerns the relationship between understanding and ability or know-how, with Hills (2016) being an influential defender of this connection.

✉ Lewis Ross  
l.ross2@lse.ac.uk

<sup>1</sup> London School of Economics and Political Science, London, UK

of intellectual disciplines like the sciences more generally.<sup>2</sup> Key to her position is the thought that a strictly factive conception of understanding is empirically inadequate. Elgin argues that it cannot account for the manifest success of scientific theorising, including the widespread role of idealisations in successful research and the role of false theories in the upwards trajectory of scientific understanding. While these considerations have long been discussed by philosophers of science engaged in the realism/antirealism debate, what is notable about their appearance in this context is how they are used to make a bold claim about the nature of an *epistemic state*—namely the claim that understanding is not factive.

Using Elgin's view as a foil, this paper shows that a strictly factive theory of understanding has resources with which to explain the essential role of false theories and idealisations in science. While the role of idealisations has been increasingly discussed in recent work, this paper aims to build on this work by offering a unified response to *both* the problem of idealisations and the role of false theories in the upwards trajectory of scientific understanding. By showing how subjects can *recover* true beliefs from false or partially inaccurate inputs, I argue that strictly factive conceptions of understanding can capture the epistemic improvements made by inquirers even if they happen to be unaware of the falsity of the view in question or the idealised status of the model or law being employed. The strategy outlined also holds promise in explaining why a factive account of understanding is compatible with the role of non-propositional (and hence not truth-apt) representations in the acquisition of understanding. In sum, strictly factive theories of understanding can readily overcome a number of their most challenging criticisms.

## 2 Theoretical Preliminaries

We use the term 'understanding' in different ways. Elgin's concern is primarily with *objectual* understanding rather than understanding propositions or understanding why something is the case.<sup>3</sup> Attributions of objectual understanding are common and familiar, taking the form: "x understands  $\phi$  where  $\phi$  is a topic, discipline, or subject matter".<sup>4</sup>

It is also important to distinguish, as Elgin does, between the '*objects*' of understanding and the '*vehicles*' of understanding.<sup>5</sup> The former is whatever the understanding is of or about—such as chess tactics, the Bronze Age, Zydeco music, and so on. The latter is whatever metaphysically comprises the understanding—such as propositions, representations, dispositions, and the like. Given that propositions are straightforwardly truth-evaluable, a natural starting point for investigating the

<sup>2</sup> See Elgin (2004, 2007, 2017). While I use Elgin's view as a foil, other helpful discussions of non-factive approaches to understanding include: Zagzebski (2001), de Regt (2015), Potochnik (2017) and Rancourt (2017).

<sup>3</sup> See, for instance, Khalifa (2013) on the relationship between different types of understanding.

<sup>4</sup> Elgin (2017: 43).

<sup>5</sup> See, also, Greco (2014) on this distinction.

factivity of understanding is to focus on cases in which the vehicles of understanding are propositional. I will return to the question of non-propositional representations at the end of the paper. A focus on propositions as the vehicles of understanding is mirrored in Elgin's way of putting the factivity view she attacks:

**Factivity:** "If 'understanding' is factive, all or most of the propositional commitments that comprise a genuine understanding are true".<sup>6</sup>

The denial of **Factivity** is a denial of two distinct approaches to the relationship between understanding and truth. Strict factive approaches claim that we can fully account for the nature of understanding by focusing *only* on factive cognitive attitudes, such as knowledge and the acceptance or belief of true propositions.<sup>7</sup> A less strict factive view, sometimes dubbed 'quasi-factivism', argues that understanding can tolerate some false commitments so long as these are peripheral, or so long as most of the commitments comprising the agent's understanding are true.<sup>8</sup> In denying **Factivity**, Elgin rejects both approaches. Rather, she claims that understanding can consist, at its core, of false cognitive attitudes—provided that these are 'true enough' for the intellectual endeavour under discussion.

To fully make sense of **Factivity**, we must bear in mind a further point: understanding has a *subject*. In other words, understanding is an epistemic state that is possessed by epistemic agents.<sup>9</sup> To illustrate what I mean, consider that a history textbook might contain many propositions about the Peloponnesian War apt to confer understanding to an interested reader. However, it is not the case that the textbook understands the Peloponnesian War. Talk of books *containing* epistemic states like knowledge or understanding is just an attenuated way of saying something like 'a reader can acquire knowledge or understanding from this book'.<sup>10</sup> So, when we discuss the vehicles of understanding (here we are concerned with propositions) concerning some object (e.g. Zydeco music) we must remember to think of these in the context of being cognitively endorsed by an epistemic agent.

A chief concern for Elgin in denying **Factivity** is to undermine a foundational view in epistemology that she dubs 'veritism', namely the idea that truth-conduciveness is the only appropriate standard of assessment for epistemic policies, practices, and their products.<sup>11</sup> Rather, Elgin wants to replace veritism with a view on which

<sup>6</sup> *Ibid.* 37.

<sup>7</sup> See, e.g., Strevens (2008, 2013), Greco (2014), Khalifa (2017), Lawler (forthcoming), and Nawar (forthcoming) for those who can be read as sympathetic to a strictly factive view. See Sliwa (2015, 2017) and Kelp (2015, 2017) for a focus on knowledge.

<sup>8</sup> For example, see Kvanvig (2003), Mizrahi (2012).

<sup>9</sup> Here, taking the simplest case, I focus on individual epistemic agents. However, recent work in social epistemology is increasingly sympathetic to the thought that *groups* can also constitute epistemic agents. For instance, Bird (2010) or Ross (2020) discuss the scientific community as an epistemic agent capable of possessing epistemic states like knowledge. There has been very little discussion of group understanding (although see Boyd 2019), but any approach on which group understanding involves the group taking on propositional commitments can, in principle, be given a treatment consistent with **Factivity**.

<sup>10</sup> We also say things like 'this book contains *an* understanding of X' when we mean something like it contains an *interpretation* of X. I won't be concerned with this type of meaning here.

<sup>11</sup> Elgin (2017: 9).

understanding rather than truth is the central aim of many of our epistemic practices, including scientific theorising, where understanding does not conform to **Factivity** but rather is a matter of being able to ‘exploit information’ and to “know how to wield it to further [one’s] cognitive (and perhaps practical) ends”.<sup>12</sup>

Before moving on, a final note about epistemic value and the factivity of understanding is needed.<sup>13</sup> Elgin quickly moves from characterising veritism in terms of the centrality of truth-conducive standards of assessment, to equating veritism with a view on which only the production of true beliefs has epistemic value. One might question this equation. For example, some have argued that we can preserve the core elements of a veritistic epistemology even if we agree that there is more to epistemic assessment than asking whether or not a given practice generates a strictly true commitment. Warenski (forthcoming) defends an ‘ecumenical veritism’, on which veritism is constituted by a plurality of truth-related norms rather than an overarching concern with true belief; Le Bihan (forthcoming) argues that Elgin’s own view, which focuses on ‘true enough’ content, is nonetheless veritistic because partial truth remains the benchmark for epistemic assessment, and Nawar (forthcoming) argues that we might preserve veritism as a thesis about the value of true belief by taking up an attitude of acceptance rather than belief towards the felicitous falsehoods identified by Elgin.

While the debate over the factivity of understanding is tightly interlinked with debates concerning veritism and the epistemic value of truth, they are independent. After all, one can argue over whether understanding is factive, quasi-factive, or non-factive without taking any stance on the epistemic value of understanding or the cognitive commitments that comprise it. In this article I focus only on the question of the factivity of understanding, leaving aside any considerations of epistemic value. The ambition of my paper is to show that we can accommodate the insights behind Elgin’s work within a *strictly* veritistic framework, eschewing both quasi-factive and non-factive approaches. Nonetheless, readers will be able to see for themselves how the conclusion I draw—viz. that understanding is necessarily comprised of true beliefs—fits into to their commitments regarding epistemic value.

### 3 Two Arguments Against Factivity

Elgin pushes two distinct lines of thought in order to undermine **Factivity**. The first appeals to the role of idealisation in science. I will outline two influential responses to the prominence of seemingly inaccurate idealisations and indicate what are by now familiar issues with these responses. The second argument against **Factivity** appeals to the improved understanding afforded by certain ‘felicitous falsehoods’: this argument cannot be dealt with by using the familiar responses to the argument

<sup>12</sup> *Ibid.* 46.

<sup>13</sup> Beyond Elgin, influential early defenders of the idea that understanding has distinctive epistemic value include: Zagzebski (2001), Kvanvig (2003), and Pritchard (2009).

from idealisation. After laying out these arguments, this paper will show how we can defuse both of them while maintaining a factive conception of understanding.

Before offering my diagnosis of how to defuse these arguments, it is worth flagging that I will not appeal to any distinction between belief and acceptance in order to do so. When developing her view, Elgin discusses possessing understanding both in virtue of having certain beliefs and *accepting* certain theories. Some have argued that closely distinguishing belief and acceptance can do important work in defending Factivity against arguments for non-factive approaches to understanding.<sup>14</sup> For example, if it is supposed that believing something that we explicitly take to be false is epistemically unacceptable, it may still be possible to *accept* the same content for theoretical purposes without it having the implication that we knowingly believe a falsehood. Here, I will offer a way to defuse Elgin's arguments that is compatible with supposing that understanding is something that we possess simply in virtue of full-blooded beliefs in the content of scientific theory. In this paper, henceforth, I will move freely between talk of belief and acceptance.

### 3.1 The Idealisation Argument

One of Elgin's arguments against **Factivity** appeals to the widespread use of *idealizations* in science.<sup>15</sup> Idealisation occurs when scientists devise models which simplify and abstract away from what they describe in order to help us theorise about and understand a target phenomenon. A standard example of idealisation is the fittingly named ideal gas law; it describes the behaviour of gases composed of molecules that lack extension and do not exhibit intermolecular attraction. Such gases cannot exist. So, the thought goes, the ideal gas law is actually false. Nonetheless, as Elgin points out, the ideal gas law is at the heart of contemporary thermodynamics—it is not an aim of science to *replace* this law because of its idealised nature, despite the fact that, strictly speaking, it fails to accurately describe gases. Hence, she concludes, “It is simply not the case that the bodies of information that constitute scientific understanding are, or that their ultimate successors can be expected to be, composed of truths, with any residual falsehoods occurring only at the periphery”.<sup>16</sup> In other words, the central and ineliminable role of idealisations in science supposedly undermines **Factivity**.

This type of argument has received attention elsewhere, including in literatures not primarily concerned with understanding. For instance, idealisations have been used as a central motivation for antirealism in the philosophy of science.<sup>17</sup> One

<sup>14</sup> See Nawar (forthcoming) and Khalifa (2017: chapter 6) for more on using the belief/acceptance distinction in the context of the present debate. For work offering a closer examination of the difference between the belief and acceptance and its relevance for understanding, see Dellsén (2017).

<sup>15</sup> For further discussion of arguments from idealisation and scientific understanding see, among others, Strevens (2008), Bokulich (2011), Mizrahi (2012), Rice (2015, 2019), Potochnik (2017), de Regt (2017), Sullivan and Khalifa (2019) Lawler (forthcoming) Nawar (forthcoming), and Frigg and Nguyen (forthcoming).

<sup>16</sup> *Ibid.* 62.

<sup>17</sup> For discussion, see, e.g., Psillos (1999), Bird (2007).

well-known response to the problem of idealisations appeals to the idea of ‘verisimilitude’. This line of response suggests that while idealisations might not be true there is a sense in which they are ‘closer to the truth’ or ‘more truthlike’ than their predecessor theoretical commitments.<sup>18</sup> In this sense, scientific theories can move closer to the truth while being composed of false propositions. Given the roots of this approach in the much-discussed philosophy of science due to Popper, there is much to be said about a verisimilitudinarian strategy to defusing the objection from idealisation.<sup>19</sup> One obvious worry concerns whether it can really deal with all of the relevant idealisations; for instance, while we might use a more complex and less idealised law to discuss the behaviour of gases, it is unclear that this would be theoretically preferable in all cases despite being closer to the truth. However, discussion of this issue in detail is orthogonal to our purposes because appealing to verisimilitude does not allow for a genuine defence of **Factivity**. Despite focusing on coming closer to the truth, the verisimilitudinarian approach still concedes that the propositions key to an improved understanding of the world are indeed false. This is because truth, at least when applied to propositions, is not a property that comes in degrees. Hence, to defend **Factivity**, it is necessary take on the more ambitious project of vindicating the full-blooded truth of the propositions comprising an agent’s understanding.

A second realist response, one that preserves a commitment to full-blooded truth, is to suggest that while the ideal gas law might be strictly speaking false it is nonetheless ‘approximately true’. So, although  $\langle p \rangle$  may be a false proposition,  $\langle \text{approximately } p \rangle$  is a true proposition. Hence, by taking this line, a friend of **Factivity** might suggest that any understanding of thermodynamics conferred by the ideal gas law consists in full-blooded truths of the form  $\langle \text{approximately } p^1; p^2; \dots p^n \rangle$  rather than in falsehoods.<sup>20</sup> This response faces a number of difficulties. The first is providing a satisfactory account of what notion of approximation we should be working with—there is now a history of failure in articulating such accounts with any degree of precision.<sup>21</sup> A further issue is whether the approximation response can really capture *all* forms of idealisation. For instance, Snell’s refraction law is an idealisation that is true less than 50% of the time. Can it be said to be approximately true? It is far from clear that this is plausible. Another problematic example comes from idealised models in population biology which (of course, falsely) represent populations as having infinitely many members. It is not obvious how we should think of the claim that any finite number is ‘approximately’ infinite.

<sup>18</sup> Niiniluoto is a modern defender of the use of verisimilitude in philosophy of science, see his (2014) for succinct discussion. See Rowbottom (2015) or Mizrahi (2017) for contemporary critique.

<sup>19</sup> See in particular Popper (1963, 1976). As has been noted elsewhere (e.g. in de Regt 2018) there are unanswered questions as to exactly how Elgin’s own view intersects with defences, such as Popper’s, of verisimilitude. Indeed, at some points Elgin’s view—appealing to ‘true enough’ falsehoods—might naturally be categorised as a verisimilitudinarian view.

<sup>20</sup> The most sophisticated version of this view is pressed by Bird (2007). Note that Bird focuses on knowledge and not merely truth. See, for example, Saatsi (2019) for discussion and critique.

<sup>21</sup> For example, Psillos (1999: 261–275).

A third response to idealisations draws on a theory of scientific explanation, the *difference-making* theory (see, in particular, Strevens 2008, 2012 and Khalifa 2017). The difference-making view argues that within any causal nexus precipitating some event *e*, only some elements of this nexus are *difference-making*. For instance, using an example from Salmon (1997), Strevens explains that both an errant baseball and the shouts of the players are part of the causal nexus that lead to a window smashing (the explanandum). However, only the former is a difference-maker in making it such that the explanandum holds (the window smashes), even if both cause the molecules in the window to vibrate. Similarly, one might argue, as Strevens (2008: chapter 8) does, that the idealisation that gases do not collide is not a difference-maker with respect to the fact that gases conform to Boyle's law.

One tactic for defusing the centrality of idealised falsehoods to scientific explanation is to deny that the inaccuracies contained within idealised models are difference-making: rather, they are *non-difference-makers*. Thus, Strevens writes:

An idealizing explanation does not incorrectly represent certain fictional factors as making a difference to the explanandum, then, but rather correctly represents certain actual factors as not making a difference. For this reason, such an explanation is superior to its veridical counterpart [...] [Strevens 2012: 456].

I have some sympathy for this theory.<sup>22</sup> Nonetheless, as Lawler (forthcoming: 7) points out, it would take further theorising to explain how this theory, which is essentially about scientific explanation, relates to the content endorsed by an epistemic subject who possesses understanding. Another important project is to evaluate accounts that aim to cleanly delineate difference-making from non-difference-making influences.<sup>23</sup> Here, I press a different type of worry: there is a second argument due to Elgin that seems to elude the non-difference-making account. It is to this argument that I now turn.

### 3.2 The Upwards Trajectory Argument

Elgin's second argument against **Factivity** starts from the idea that understanding is a gradable notion—it comes in degrees. Elgin uses this platitude about understanding in order to generate trouble for the thought that only true propositions can comprise our scientific understanding. Specifically, she claims that it leaves us unable to account for cases in which we endorse certain 'felicitous falsehoods' and thereby move upwards along the trajectory from ignorance towards a better understanding of some subject-matter. This is endemic within scientific education. Consider the following simple case:

<sup>22</sup> Although see Rice (2015) for criticism. His critique appeals to optimality models within biology, an example to which I return in S.6.

<sup>23</sup> For summary of Strevens' *kairitic* model, see Strevens (2012).

An eight-year-old's understanding of human evolution might include as a central strand the proposition that human beings descended from apes. A more sophisticated understanding has it that human beings and the other great apes descended from a common hominid ancestor who was not, strictly speaking, an ape. The child's opinion displays some grasp of evolution. It is clearly cognitively better than the belief that humans evolved from butterflies. But it is not strictly true. [...] [T]he pattern exhibited in this case is endemic to scientific education. We typically begin with rough characterizations that properly orient us toward the phenomena, and then refine the characterizations as our understanding of the science advances. [Elgin 2017: 59].

As Elgin suggests, it is natural to suppose that the child acquires a better understanding when they come to believe the false proposition that humans descended from apes. Although learning such a proposition might not be as good as *knowing* that we instead descended from a common hominid ancestor, the falsehood nonetheless contributes to one's understanding of evolution. Hence, it seems, factive accounts of understanding are unable to explain this improvement.<sup>24</sup> We can call these scenarios where felicitous falsehoods seem to improve our understanding of some subject matter '*Elgin Cases*'.

It seems to be a non-starter to suppose that the falsehood 'humans descended from apes' is a *non-difference-maker* in the proffered explanation, just as the shouts of a collection of baseball players is not a difference-maker vis-à-vis the window smashing. If the proffered explanation were true, then clearly the false proposition ('humans are descended from apes') would make all the causal difference the explanandum ('why humans are thus and so'). Moreover, in summarising the difference-making theory, Strevens suggests that the irrelevance of false non-difference-makers is secured by the fact that scientists have the facility to properly interpret idealised models correctly so as to focus only on the relevant and true causal factors.<sup>25</sup> But this will not do here: an eight-year-old is not a scientist who is applying their knowledge of how to interpret the explanations offered to them.

Nor can the defender of **Factivity** respond to Elgin Cases by appealing to the 'approximation' idea that was available in response to the argument from idealisation. Or, at very least, that response will not work across the board. I want to bypass any local debate as to whether it is approximately true that humans are descended from apes because other examples more clearly elude the approximation strategy. To see this, let's consider another example discussed by Elgin. While our original example concerned scientific education, Elgin's objection against factive theories of understanding in fact ramifies to cases in which falsehoods are not merely used

<sup>24</sup> A different way to respond to these arguments, short of agreeing that false theories afford genuine scientific understanding, is to argue the apparently progressive nature of false theories is explained by the fact that they afford agents with a sort of procedural understanding (e.g. of how to conduct certain experiments) rather than explanatory understanding, or with a sort of proto-understanding that is a mere precursor to genuine scientific understanding (see Khalifa 2017: chapter 6 for a discussion in relation to phlogiston theory).

<sup>25</sup> E.g. see Strevens (2012: 456).



as a crutch for educative purposes, but actually represent the cutting-edge of scientific belief. The example is that scientists improved their understanding of the cosmos by adopting a series of false yet comparatively better cosmological theories: from Copernicus' heliocentric model on which the earth has a circular orbit around the sun; to Kepler's model on which the orbit is elliptical; to Newton's model that accounted for gravitational attraction and posited an irregular elliptical orbit; and so on to the present day (see Elgin: 2017: 60). Now, one might still suppose that each of these theories better approximates the truth than the previous. However, this response cannot plausibly be maintained for all predecessor theories. Consider a pre-Copernican transition; the adoption of a geocentric model compared to (let's say) mythical cosmologies that explained the movement of the celestial bodies in terms of divine activity. The adoption of a geocentric model might reasonably be thought to have constituted progress on the trajectory from ignorance towards a sophisticated understanding of the cosmos. Yet, I take it, we cannot plausibly say that it is even approximately true that the sun revolves around a static earth.

What we need is a unified way of responding to these two arguments against strictly factive conceptions of understanding. The strategy I will propose is a simple one: I argue that, on close inspection, the cases of improved understanding Elgin appeals to can in fact be explained by attributing new true beliefs to the relevant subject. In the next section, I lay the foundations for this response.

#### 4 Falsehood and the Recovery of Truth

Elgin Cases both compel and require us to explain the following concession: falsehoods *can* play a valuable role in scientific education. While it is somewhat counterintuitive to suppose false testimony to be educationally beneficial, brief reflection shows Elgin's insight about scientific and indeed other types of education to be both correct and important—we routinely and often appropriately instruct novices by telling them things that are not strictly true.

I suggest that we can accommodate this insight within a fundamentally veritistic framework. Explaining why is an important preparatory step that will enable us to later defend a strictly factive theory of understanding. Falsehoods can be useful in the context of education for the following reason: directing someone's attention to a *tactically chosen falsehood* is often a simple way to *enable one's interlocutor to recover true information*. This educational strategy draws on a very general feature of human communication, the fact that listeners are able to recover a variety of different contents from a piece of information. This is best appreciated with an example.

Imagine, for instance, that a bartender in London is making small-talk with a tall blonde stranger who has just walked into the bar. The bartender asks the man 'Where you are from?' and they reply that they are from Helsinki. Let's now suppose that their response was false—the tall blonde stranger is actually from Espoo, a nearby city in southern Finland that almost forms a conurbation with Helsinki. They have asserted a falsehood. However, assuming the bartender has the ordinary set of associations with Helsinki, there are good epistemic and pro-social reasons for the tall

blonde stranger to assert the false proposition in this situation.<sup>26</sup> Namely, the stranger knows that by stating the false proposition that they are from Helsinki, they will thereby convey to the bartender a variety of germane true beliefs.<sup>27</sup> These include: He is Finnish; He is from Northern Europe; He is accustomed to cold winters; and so forth.<sup>28</sup> Crucially, were the tall blonde stranger to have *truly* asserted that they are from Espoo, the bartender, probably not knowing the first thing about Espoo, would have acquired *less in the way of true beliefs* about the tall blonde stranger.

Hence, we can explain the epistemic value of certain tactical falsehoods within a veritistic framework. This account of the utility of employing tactically chosen falsehoods applies equally to educational contexts. Sometimes in order to convey important true beliefs in a way that is economic and easy to digest, it can be effective to do so by focusing the attention of your interlocutor on a false proposition tactically chosen to convey these truths. Focusing right away on the truth can be epistemically deleterious in the sense not only of inviting confusion and unwanted questions, but also in the sense of conveying less in the way of true belief. For instance, telling an eight-year-old that humans evolved from a ‘common hominid ancestor’ is a likely example of true testimony that conveys *less in the way of true belief* than a tactical falsehood. When exactly we should resort to a tactical falsehood rather than the strict truth depends on the situation—particularly on the cognitive abilities of one’s interlocutor, and on how important precision is in the relevant context—but discussing these parameters is not my purpose here. Rather, the main lesson of this section is that focusing on a falsehood can involve (and even be an effective way of) learning a variety of true beliefs. This is so even while preserving the thought that transmitting truth is the ultimate aim of a given communicative exchange, and that the acquisition of truth on the part of one’s interlocutor is the evaluative benchmark for the success or otherwise of that exchange. By recognising this, we can uphold Elgin’s insight about the role of falsehood in education while maintaining the veritistic principle that transmission of truth is still the central (and only) epistemic aim.

The focus on *recovery* of true content has important affinities with another strategy suggested by recent literature. For instance, it has been argued scientists knowingly use false models to enable them to purposefully *extract* accurate information about the phenomenon under investigation (see, in particular, Lawler forthcoming; and see Rice forthcoming for the connection with scientific realism). However, one feature of the idea outlined here is that the recovery of true information from false testimony can sometimes be both *inadvertent* and relatively *automatic*. In this sense, the recovery of true information is entirely consistent with it being the case that

<sup>26</sup> Of course, people might not have the associations we (reasonably) expect them to have regarding certain pieces of information. Nothing in my discussion turns on this, so long as we grant that there are situations in which we can reasonably expect people to have the right associations.

<sup>27</sup> The example still illustrates this point, even if there might be *other* locutions that convey even more true beliefs such as ‘I am from a city in the neighbourhood of Helsinki’.

<sup>28</sup> This point is not wedded to but follows from various popular views of the nature of belief. For instance, provided the bartender knows that Helsinki is in Finland, attributing the belief <the stranger is from Finland> to the bartender would be warranted both by looking at what the bartender would be *disposed* to endorse and as an *interpretation* of their outward behaviour (e.g. future conversational moves).

the endorsement of what Elgin calls a felicitous falsehood is entirely unwitting—an agent can unknowingly learn something false, but still thereby come to acquire a host of true beliefs about the subject of inquiry. It is precisely this fact that makes felicitous falsehoods so powerful within the context of education.

Having taken the preparatory step of explaining how it is possible to recover true content from falsehood, we can now turn to defuse the arguments against a strictly factive theory of understanding.

## 5 Defusing the Upwards Trajectory Argument

One argument against **Factivity** was that only a non-factive conception of understanding can explain Elgin Cases in which endorsing falsehoods moves us along the trajectory from ignorance towards sophisticated understanding.

To begin to defuse this argument, we can start by reconsidering Elgin's example regarding an understanding of evolution. In this case, the false proposition identified as bringing about an improved understanding of evolution was:

(F) Human beings are descended from apes.

Let's immediately grant that the child believes (F) and their understanding of evolution increased as a result whatever learning experience caused them to endorse (F). Although Elgin's example has us focus on one particular proposition, we noted in the previous section that coming to endorse a (strictly false) proposition can cause an agent to thereby recover a number of associated true propositions. Indeed, it is almost inconceivable for an agent to have a *solitary* belief about a given subject-matter.<sup>29</sup> With this in mind we can then provide an alternative diagnosis, one that preserves **Factivity**, for the improved understanding in Elgin Cases. Specifically, I think the improved understanding afforded by Elgin Cases can be fully captured by appealing to true beliefs acquired by the agent in question.

Some natural examples of the other contents that one would typically come to endorse as a result of accepting (F) are the following:

- (T1) Humans have non-human ancestors.
- (T2) Humans descended from hairy creatures that were sometimes bipedal.
- (T3) Humans and apes have a lot in common.
- (T4) Humans are the way they are due to gradual change over time.

The thing to notice about this list of contents is (at least as generalisations or approximations) they are all *true*. My claim is that it is *only* in virtue of thereby endorsing

<sup>29</sup> A helpful quote from Braddon-Mitchell and Chalmers illustrates:

What would it be like to believe that there's milk in the refrigerator, and nothing else? It seems as impossible as having money without the social and economic circumstances that give sense to something being money. To believe that there is milk in the refrigerator, you have to have enough by way of belief to count as understanding what milk is, what a refrigerator is, and what it is for one thing to be inside another. It takes a lot of belief to be any amount of belief (Braddon-Mitchell and Jackson 2007: 196).

these or similarly true contents does an agent come to increase their understanding of evolution. The child's understanding of evolution does not consist in their belief in the false proposition (F). Rather, this proposition is only instrumentally valuable insofar as it leads them to endorse true contents. We can support this conclusion with two distinct lines of thought.

## 5.1 Tracking

The first consideration in favour of an entirely veritistic explanation of Elgin Cases is that *attributions of understanding track the true contents* endorsed by subjects, even when subjects share belief in the very same false proposition.

We can tease this out by noting that among children who each believe, falsely, that human beings are descended from apes, there can be variation in which associated true contents they also endorse. This follows from the observation that, for many propositions  $p$ , it is impossible to present a list of contents that we *must* endorse in order to count as believing  $p$ . There are different explanations for this difficulty. One follows from growing scepticism among philosophers and psychologists about the 'classical' theory of concepts on which concept-possession requires categorising the objects of belief in accordance with a discoverable set of necessary and sufficient conditions.<sup>30</sup> A second would stem from the truth of even a modest version of semantic externalism. As Burge (1979) argues, it seems possible to have beliefs about some phenomenon—in his example arthritis—while lacking (or even having misconceptions about) certain core aspects of that phenomenon. Which explanation we accept is unimportant for our purposes; what matters is that we can readily imagine someone who believes (F) while lacking belief in one of (T1)–(T4). For instance, we can imagine a young child who lacks belief in (T4) because although they believe that humans are descended from apes, they labour under the misconception that this change happened all at once such that one evening there were apes and then—POOF!—the next morning there were humans.

My claim is that our attributions of understanding track the true contents of what we believe.<sup>31</sup> So, imagine two children, both of whom falsely believe that humans are descended from apes, but one believes that this was the result of gradual change while the other thinks it happened instantaneously. Clearly, the first child who believes that there was gradual change has better understanding of evolution than the second who believes the change happened all at once. This isn't merely an intuition, but follows naturally from two ideas that have widely been taken to be diagnostic if not constitutive of superior understanding: (1) privileging more fine-grained correct explanations for why things are the way they are (e.g. see Sliwa 2017), and (2) having improved abilities to reason and draw inferences regarding the subject of one's understanding (e.g. see Hills 2016). In this case, the first child privileges a

<sup>30</sup> See Murphy (2002: chapter 2) for an overview and discussion of empirical research.

<sup>31</sup> To be clear: we will need to qualify this further (e.g. by focusing on causal beliefs, or beliefs that afford us with certain abilities), it is not the case that *any* true belief on a given subject-matter will increase our understanding of it.

more fine-grained correct explanation for why humans are the way they are (namely, due to gradual change over time), and they will also be better-placed to make accurate inferences about human evolution (e.g. about why we see different evolutionary changes happening at different points in human history). In sum, attributions of understanding seem to track the true contents of what is endorsed even if both agents share belief in the very same ‘felicitous falsehood’.

## 5.2 Plateauing

The *tracking* argument suggests that our judgements about an agent’s degree of understanding varies with the true contents of what they believe. Although instructive this does not vindicate **Activity**, for it may be the case that attributions of understanding also increase with the false contents of belief. For example, perhaps it is the case that (1) false beliefs by themselves can also supply a modicum of understanding, or (2) false beliefs can supply understanding in conjunction with true beliefs. I argue against these positions with a second argument, the *plateauing* argument. The idea is as follows: false commitments only give the appearance of constituting our understanding, in reality understanding increases just in virtue of the fact that that coming to accept these commitments *causes* the agent to endorse additional true commitments—increases in understanding occur *only if* the agent comes to endorse additional true content. In other words, our understanding plateaus with what is truly believed. This suggests that what *constitutes* an agent’s increased understanding is true content that they endorse, with false beliefs playing a merely facilitative role, even if the acquisition of the relevant false beliefs is indispensable to the causal explanation of how the agent came to acquire understanding in the first place.<sup>32</sup>

To test this claim, we should consider our judgements about the following type of case. Imagine two children who each believe all of the true contents noted in (T1)–(T4), but do *not* believe that humans were descended from apes. (Let’s suppose that they both start as epistemic duplicates, with identical dispositions and abilities. This is of course an artificiality, but not a vicious one—we could just as well imagine the same agent in two different counterfactual scenarios.) Then imagine that one of them is told that humans are descended from apes, comes to believe this testimony, but does not learn or infer anything else that could be characterised as true that the other child does not already believe (even as a suitably qualified approximation or generalisation). That is, we hold the beliefs of the two children entirely fixed apart from attributing an additional false belief to one of them. Does the child improve their understanding of evolution in virtue of accepting (F)? It is hard to see why this would be the case, as the only thing to distinguish them is a misleading

<sup>32</sup> For instance, the ‘bartender case’ from S.4 is one in which the acquisition of a false belief (‘the stranger is from Helsinki’) was a causal prerequisite for the recovery of associated true beliefs. My view is that any understanding acquired in this case is nevertheless *constituted* only by the true beliefs, even if the belief in a falsehood is central to the causal story of how the agent came to acquire a degree of understanding in the first place. Thanks to an anonymous referee for pressing me to clarify here.

commitment. A natural thought might be that learning some false belief could sometimes enable us to draw inferences yielding additional true beliefs, thus increasing our understanding of some phenomena. Does this mean that the falsehood might (partly) comprise our understanding of that phenomena? Although tempting, this idea doesn't work here. As Sliwa (2017: 524–525) explains, it is important to distinguish between the *capacity to understand* and *instances of understanding*. One can have the capacity to understand something without actually (yet) understanding it. In the case in which a false belief puts someone in a position to infer truths that would contribute to their understanding, they have the capacity to increase their understanding but have not (yet) done so. Only when they infer the truth will they exercise their capacity to understand and actually increase their understanding. And this, of course, is no problem for **Factivity**, for it is thus the acquisition of a *true* belief that explains the increase in understanding. Hence, while false beliefs might play a causal-explanatory role in why an agent possesses understanding of a given phenomenon, it is true beliefs that fully explain why they understand to the degree that they do.

Thus, we see that the addition of (F) does nothing to augment an agent's understanding once we are careful to separate it out from all of the *true* contents that we can learn by focusing on (F). So, not only do understanding attributions *track* what is truly believed, increases in understanding seem to *plateau* with what is truly believed too. As such, there is no reason to suppose that Elgin Cases are any threat to **Factivity**, the idea that understanding is comprised of true contents we endorse.

I have not settled the question of whether having a false belief could act a negative modifier which *lowers* one's degree of understanding. To answer this question, we would need to compare cases where an agent is agnostic about some proposition against cases where an agent has an additional false belief about the phenomena under investigation. While it seems plausible that such false beliefs could lower one's understanding—e.g. if a child thought that humans were specifically descended not just from apes, but from orangutans, or that humans are descended from capuchin monkeys, which they falsely supposed to be an ape—my suspicion is that these false beliefs would only inhibit understanding when they precluded an agent from believing some true content we would otherwise expect them to have. Understanding, on this view, would be fully determined by the true contents of what one believes, even though additional false beliefs might serve to prevent an agent from believing certain true contents. This is a preliminary sketch of a response to this issue; the negative influence of false beliefs on attributions of understanding is an important question for future research on factive theories of understanding.<sup>33</sup>

### 5.3 Approximation Revisited

Before moving on, it is worthwhile to briefly observe that the broader strategy of teasing out 'strictly' true contents of what agents believe in Elgin Cases succeeds

<sup>33</sup> My thanks to anonymous reviewers for pressing me on this issue.

where appealing to the idea of approximate truth fails. The earlier example of a problematic case for the approximate truth response was that of moving to a geocentric model of the cosmos in place of a mythological conception. It is not ‘approximately true’ to suppose that the sun revolves around the earth; precisely the opposite is true. However, in coming to believe the falsehood that the sun revolves around the earth, we can suppose that early astronomers did also come to believe the following sorts of *true* contents:

- (T5) Many celestial bodies have regular orbits.
- (T6) Many celestial bodies move according to mechanistic laws.
- (T7) Many celestial bodies are not moved by divine willing.

Again, the factivist about understanding can claim that, to the extent we credit Ptolemaic astronomers with an improved understanding of the cosmos, this was only in virtue of their coming to learn these truths. This is the case notwithstanding their ignorance about the fact that the sun does not revolve around the earth. Moreover, to the extent that we discovered that such astronomers did not endorse contents such as (T5)–(T7), we would accordingly attribute them a worse understanding of the cosmos than if they had. Again, our attributions of understanding track and are explained by the true contents of what is believed.

## 6 Defusing the Idealisation Argument

The remaining argument against **Factivity** relies on the prominent role of idealisations in providing scientific understanding. According to this argument, a factive conception of understanding cannot account for the essential role of models and theoretical posits that are false because they simplify and abstract from reality.

Just as with my response to the upwards trajectory argument, I will argue that focusing on the recovery of true belief from inaccurate inputs allows us to defuse this argument. My analysis will be broadly complementary to the views of Lawler (forthcoming) and Rice (forthcoming), who argue that the central role idealisations in scientific theorising is compatible with strictly factive views of understanding insofar as scientists who use them consciously *extract* true information from them. However, when defusing the upwards trajectory argument, I argued that the recovery of true contents from false theories—and hence the acquisition of understanding—can be compatible with ignorance that one is considering a false theory. This is why, for example, children can recover true contents from theories that they do not realise are false. In a similar vein I will here argue, at least in some cases, that something similar holds for idealisations: they permit the *recovery* of true content even if an agent is unaware of the idealised nature of the model or theory. Thus, I will vindicate the extraction view by suggesting that it is a species of a general fact about scientific theorising: the arguments, models, and theories we use provide understanding to the extent that they allow an agent to recover true content from them, whether wittingly or unwittingly.

To begin to respond to the idealisation argument, recall one of platitudes we observed at the outset—namely, that understanding is an epistemic state that requires a *subject*. Just like beliefs require a believer, and knowledge a knower, understanding presupposes an agent doing the understanding. This is why, for instance, a textbook ‘provides understanding’ only in the sense that it elicits certain epistemic states in the agents making use of it. Keeping this in mind explains the following rather obvious point: when it comes to thinking about the role of prominent idealisations in scientific practice—such as the ideal gas law, or models in population biology—there can be *variation* in how much is understood by the individual using the idealised law or model. This is a natural corollary to the truism Elgin earlier relied upon in making the upwards trajectory argument, namely that understanding comes in degrees. Different scientists using idealised laws and models can understand their subject-matter to different degrees.

The reason this matters is that, since understanding is something possessed by epistemic agents, we must interpret the argument from idealisation against **Factivity** accordingly. In this vein, the argument cannot succeed simply by noting that the idealised models or laws themselves contain falsehoods. Rather, the focus must be on how idealisations *elicit* understanding in those who use them. And, in order to plausibly undermine **Factivity**, the argument from idealisation must say: it is *not* possible to explain the understanding elicited by idealised models and laws only by appealing to true beliefs acquired by those who use them. Put in these terms, so I claim, the argument from idealisations against **Factivity** will turn out less compelling than it initially sounded.

As has been pressed extensively by Lawler (forthcoming) and Rice (forthcoming), when using idealised models and laws, scientists are typically aware of them being idealisations. For example, when using the ideal gas law, scientists are aware that the real-world gas they are theorising about is not composed of molecules which lack extension. The emphasis on this knowledge chimes with comments in Strevens (2012: 456), who suggests that the inessential role played by falsehoods in scientific idealisation is partly explained by the fact that scientists “know the right way to read idealized models”. These points support the following schematic observations: (1) firstly that idealisations, although false, do not invariably elicit false beliefs in those who use them, and (2) that our judgements about understanding vary with what is in fact endorsed by the agent using the particular idealisation. However, while I agree with these schematic observations, I think that it is possible to recover a degree of understanding from an idealised model—just as in the case of scientific education—even if one is unaware of its status as an idealisation.<sup>34</sup>

As the argument from idealisation is primarily driven by examples, with much of the debate turning on the proper interpretation of particular applications of certain models, we should consider how scientists use particular idealisations in context.

<sup>34</sup> Not everyone agrees. For instance, Greco (2014: 297) suggests that a scientist who was unaware of a model’s status as an idealisation would fail to gain understanding of the phenomenon it describes but rather only understanding of the representation itself (in the case he discusses, the ideal gas law describing gases).



There are various noted examples of false idealisations or models conferring understanding. The ideal gas law is one, but this example has been widely discussed.<sup>35</sup> Another type of example, one that has figured in the debate concerning the nature of scientific understanding, is the use of models in population biology. Indeed, these have been used by Rice and Lawler to test the plausibility of factive views of understanding. For example, one closely analysed case is the use of optimality models in biology, such as the attempt to work out the optimal copulation time for male dung flies visiting multiple piles of dung in order to mate.<sup>36</sup> However for the sake of any pre-breakfast readers I shall focus on a different optimality model.<sup>37</sup> To do so, we will need to talk about crows and whelks.

## 6.1 Crows and Whelks

There is colony of Northwestern Crow on Mandarte Island in British Columbia. These crows feed on whelks, a type of mollusc that lives in a hard shell. The crows open the whelks by dropping them onto a rocky beach; they only select whelks that are above a certain size; they almost always drop the whelks from a height of around 5 m; and the crows don't tend to give up if a particular whelk stubbornly refuses to break after a few drops.<sup>38</sup> Upon hearing these facts, we—or a zoologist—might be interested in better understanding why the crows forage in this way. This is where optimality models can be useful.

Optimality models in biology help us to understand why a given population possesses a particular trait, by showing that the trait in question maximises evolutionary fitness in light of certain constraints. When it comes to foraging, the relevant trade-offs are calorific; what behaviours strike the right balance between energy expenditure and calorie-acquisition? The predicament facing our crows is how to achieve the right balance between energy gained from eating denuded molluscs and energy expended in upwards flight. Zach (1978, 1979) provides an optimality model demonstrating that the crows adopt an optimal strategy: focusing on large whelks (which provide more calories and break much more easily than smaller whelks), dropping them from around 5 m (which provides the best trade-off between likelihood of breakage and calories expended), and being persistent in continuously dropping their chosen whelk (because each successive whelk drop is about as likely to succeed as taking a new whelk). Hence, the crows more or less optimise calorific gains when whelk foraging.

Optimality models involve artificial idealisations and simplifications that render them inaccurate with respect to the actual causal mechanisms which led to the

---

<sup>35</sup> See, for example, Strevens (2008), Potochnik (2017), Khalifa (2017), Rice (2019), Sullivan and Khalifa (2019), Lawler (forthcoming), Nawar (forthcoming). For a recent defence of a reading of the ideal gas law that supports a non-factive theory of understanding, see Doyle et al. (2018).

<sup>36</sup> See Parker (1978) for the empirical data and Sober (2000) or Rice (2015) for philosophical discussion.

<sup>37</sup> See Lawler (forthcoming) for a more abstract discussion of strictly factive views in relation to optimality models.

<sup>38</sup> Zach (1978, 1979) provides these empirical observations.

evolution of a given trait within a target population.<sup>39,40</sup> With respect to our colony of crows, the model used is simplified in various respects. For example, when working out the calorific expenditure used in flight, the model falsely supposes that all crow flight is *horizontal* by adopting the simplifying supposition that the higher calorific costs of ascending flight and the lower calorific costs of descending flight will roughly cancel each other out. So, in order to elide complicated calculations, the model works on the basis of supposing that *flying higher* uses more calories only because it involves *flying for longer*. This obviously isn't entirely accurate, upward flight is more strenuous beyond simply extending the period of flight. Moreover, the calorific expenditure of horizontal flight in the model is calculated using a *constant base rate*; this strips out real world influences like favourable or adverse wind conditions or physiological differences that will change the actual level of calorific expenditure for a given crow. This isn't entirely accurate either—there is no single base rate that accurately captures how many calories every single crow uses when flying over a given time period. In short, the model used does not accurately represent all of the actual causal mechanisms influencing the development of the crows' foraging behaviour. Rather, it simplifies and omits various factors for theoretical ease. However, clearly such models *are* useful in helping us better understand the crows' foraging behaviour; by illustrating different types of trade-offs they face, the model helps us grasp why certain behavioural strategies are apt to be selected for.

Although the optimality model *itself* contains false idealisations, we must remember that to determine whether or not **Factivity** is in trouble we must look at the epistemic state it elicits in those using the model. So, is there any reason to suppose that the *understanding elicited by those using this model* must contain false beliefs? The answer, I think, is negative. Rather, we can readily explain the understanding elicited by such models by appealing to true beliefs. When we think about what is useful about Zach's optimality model, I suggest, we find that it is the fact that it justifies us in adopting the following type of beliefs about the Mandarte crow:

<sup>39</sup> Rice (2015: 600) identifies six different ways in which optimality models tend to idealise:

- (1) The model's mathematical curves, equations, or payoff structures are often idealized when compared to the causal processes within the target system(s).
- (2) Idealized strategy sets are intended to capture the relevant alternatives rather than strategies actually causally interacting within a population.
- (3) The models' optimization assumptions do not accurately represent a causal mechanism in the system, but only captures the general optimizing tendency of the system in the long run.
- (4) There are idealizations regarding causal mechanisms of inheritance.
- (5) It is assumed that selection pressures do not change over time.
- (6) Infinite population size is assumed to allow for the use of various laws of large numbers in deducing the target explanandum.

<sup>40</sup> Our concern here is with factive versus non-factive theories of understanding. However it is worth noting that philosophers of science have also been interested in the idealised nature of optimality models for what they tell us about a related yet orthogonal debate: on what the nature of scientific explanation is, e.g. whether a scientific explanation must accurately represent the causes of the explanandum. (For instance, see Cartwright 1983 or Bokulich 2011, 2016 for work on idealisations and scientific explanation). There are interesting connections between debates about the nature of understanding and the nature of scientific explanation, but here I only discuss the former question. For a broader perspective on both of these issues, see book-length treatments in Strevens (2008) and de Regt (2017).

- (T1) Selecting large whelks is an effective foraging strategy because those contain more calories and are more likely to break than smaller whelks.
- (T2) The disposition to drop whelks from ~5 m strikes a good balance between calorific expenditure and likelihood of opening the whelk.
- (T3) Persistently dropping the same whelk until it breaks is at least as good a strategy as selecting a new whelk.
- (T4) Natural selection will tend to favour crows which optimise calorific gains when foraging over those which are profligate with their energy.

All of these beliefs (at least as generalisations or approximations) are true, and they are central to the increased understanding about optimal foraging strategies gained from the use of the optimality model. (T1)–(T4) are the types of belief that anyone consulting Zach's model would acquire about the Mandarte crow. So far, this corroborates the idea that we can explain the understanding gained from an idealised model simply by focusing on what is purposively extracted from the model.

Although the model used to demonstrate why these beliefs are well-founded is simplified in various respects, these false simplifications are inessential to the increased understanding the model affords us with. Indeed, it is surely possible to recover a degree of understanding from such optimality models even if the observer failed to realise that the models were idealised in the respects explained above. Consider the following beliefs about the simplified false aspects of the optimality model:

- (F1) All crow flight is horizontal.
- (F2) All crows have exactly the same metabolic rate when flying.

These false beliefs are not at the heart of why Zach's optimality model offers us an increased understanding of crow foraging. Rather, these simplifications are just ways to more conveniently construct the optimality model which acts as a tool for eliciting better understanding in those who consult it. Suppose that a reader had simply not noticed that the calculations in Zach's model involved idealised assumptions about horizontal versus vertical flight or the uniform metabolic rate of crows. Indeed, I suspect that many casual readers would not immediately notice this feature of the model. Would this preclude them from understanding why the crow's chosen foraging strategy optimises calorific gains? I think that it would make very little difference, for they would still recover the relevant true contents from the model, such as those enumerated in (T1)–(T4). In this sense, I suggest that while such an unwitting reader would to some extent misunderstand the nature of the model, they would in fact acquire understanding of the phenomenon it represents. Again, as with my diagnosis of the upward trajectory argument, I suggest that we would only credit such an unwitting reader understanding *to the extent* that they acquired true beliefs from the idealisation. As such, any false belief (or, more realistically, agnosticism) regarding (F1)–(F2) would not constitute their understanding, even if they happened to be ignorant in this way.

To sum up: while Elgin is right in claiming that strictly false idealisations are extremely useful in scientific theorising, I have suggested that their usefulness consists in being convenient tools for eliciting true beliefs that facilitate an

understanding of their objects. As such, by focusing on how idealisations elicit understanding in those who use them and not just on the content of the idealised law or model itself, we can accept Elgin's insight about their usefulness while denying that it creates any pressure to accept a non-factive theory of understanding. Notably, idealisations can serve their purpose, at least in some cases and at least to some extent, even if an agent is not aware of their idealised status. Of course, here I have only discussed one such model. As Sullivan and Khalifa (2019: 679) concede in their critique of those who use the ideal gas law to attack **Factivity**, opponents can still maintain that it will be *possible* to find further examples that do in fact support non-factive views of understanding. However, while it is correct to say that the debate must be conducted case-by-case, I think that recent work has done enough to put on the defensive those who use the argument from idealisation to undermine strictly factive views of understanding.

## 6.2 Further Applications

The strategy outlined in this paper to defend strictly factive views of understanding against cases in which falsehoods seemingly play an ineliminable role in successful theorising has general application. By clearly distinguishing what the epistemic *subject* believes from the *vehicle* (e.g. a book, a model, a string of testimony) that delivers understanding, we can also defend factive views against a further charge: that valorising the importance of true propositions leaves us unable to account for the understanding provided by non-propositional representations which are not truth apt. Moving from scientific to historical understanding, Elgin (2017: 103) asks: "Should we deny that works of art afford historical understanding because they are not verbal?" She answers her own question in the negative as follows:

There is, as far as I can see, no reason to privilege the verbal over other modes of symbolization. And if we do, we exclude not just prints, monuments, and documentary films, but also diagrams, charts, and maps. To restrict historical understanding to that which is captured in a language would be costly. [Elgin 2017: 103].

While I can only provide a thumbnail sketch here, it is easy to see how distinguishing between the non-propositional and therefore not truth-apt content of a representation—like a diagram or map—can be separated from the propositional and therefore truth-apt content that an agent recovers from considering these representations. On the factive view defended in this paper, we would seek to explain the understanding afforded by such representations by identifying true propositions recovered by the agent using the representation. Thus, while we can agree with Elgin that non-propositional representations play a crucial role in affording us with understanding, we do not need to suppose that this creates any pressure to deny a factive account of understanding.

A further challenge, relating to the idea of non-propositional understanding, is due to Lipton (2009). He claim that we can derive 'inarticulate' understanding from using models such as an orrery—for example, that we can derive understanding

why the planets exhibit a certain type of motion without being able to articulate this understanding in propositional terms. I lack space to fully discuss this challenge here, but it is worth noting that a strictly factive view of understanding might respond by appealing to fact that beliefs do not necessarily need to be articulable to count as beliefs. For instance, there are reasons to suppose that women living in more benighted times both were capable of knowing that sexual harassment was wrong and have some grasp of why it was wrong, despite lacking the conceptual framework required to clearly articulate these beliefs.<sup>41</sup> It may be the case that, similarly, we can have beliefs that contribute to our understanding of a phenomenon even in lieu of possessing a conceptual framework needed to articulate them. Further work would do well to explore the conditions for attributing beliefs even in the absence of articulability to answer Lipton's challenge within a strictly factive conception of understanding.<sup>42</sup>

## 7 Conclusion

The question of whether or not understanding is factive has been a key faultline in debates about the nature of understanding, with a central motivation for non-factive views being the thought that only they can account for the manifest success of scientific theorising in increasing our understanding of the world. This paper, taking the influential arguments of Catherine Elgin as a foil, suggested that a factive conception of understanding does in fact have resources with which to overcome two prominent challenges it faces: firstly, the role of tactical falsehoods in enabling us to move along the trajectory from ignorance towards a more sophisticated understanding of various phenomena, and secondly, the idea that a factive view cannot explain the role of idealised models in eliciting scientific understanding. Closer inspection of the relevant cases suggested that there was no need to appeal to false beliefs in order to explain the increase in understanding. Focusing on a tactical falsehood is often an effective way of recovering a variety of true beliefs, even when the agent is unaware that they are considering something false or inaccurate. We can readily understand idealised laws and models as tools for eliciting true beliefs in those who use them rather than as constituting their understanding. In sum, we can agree with the insight that falsehoods often play an important role in successful scientific theorising while maintaining a strictly factive conception of understanding.<sup>43</sup>

---

<sup>41</sup> See Srinivasan (2020) for a seminal paper on this topic.

<sup>42</sup> Strevens (2013) provides further discussion of Lipton's example, and Khalifa (2013) discusses Lipton's view and the relation between understanding and explanation.

<sup>43</sup> The anonymous reviewers for this manuscript were especially diligent and deserve special thanks. I am also grateful to Jessica Brown, Miguel Egler, Alison Hills, Matt McGrath, Janis Schule, Justin Snedegar, and other friends and colleagues who provided advice, conversation and support that improved this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bird, A. (2007). What is scientific progress? *Noûs*, 41(1), 64–89.
- Bird, A. (2010). Social knowing: The social sense of “scientific knowledge.” *Philosophical Perspectives*, 24(1), 23–56.
- Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180(1), 33–45.
- Bokulich, A. (2016). Fiction as a vehicle for truth: Moving beyond the ontic conception. *The Monist*, 99(3), 260–279.
- Boyd, K. (forthcoming). Group understanding. *Synthese*, 1–22. <https://doi.org/10.1007/s11229-019-02492-3>.
- Braddon-Mitchell, D., & Jackson, F. (2007). *Philosophy of mind and cognition: An introduction*. Malden, MA: Blackwell.
- Burge, T. (1979). Individualism and the mental. In French, Uehling, and Wettstein (Eds.), *Midwest studies in philosophy*, IV (pp. 73–121). Minneapolis: University of Minnesota Press.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Oxford University Press.
- de Regt, H. W. (2015). Scientific understanding: Truth or dare? *Synthese*, 192(12), 3781–3797.
- de Regt, H. W. (2017). *Understanding scientific understanding*. Oxford: Oxford University Press.
- de Regt, H. W. (2018). Review of true enough. *Notre Dame Philosophical Reviews*. <https://ndpr.nd.edu/news/true-enough/>.
- Dellsén, F. (2017). Understanding without justification or belief. *Ratio*, 30(3), 239–254.
- Doyle, Y., Egan, S., Graham, N., & Khalifa, K. (2018). Non-Factive understanding: A statement and defense. *Journal for General Philosophy of Science*. <https://doi.org/10.1007/s10838-019-09469-3>.
- Elgin, C. (2004). True enough. *Philosophical Issues*, 14(1), 113–131.
- Elgin, C. (2007). Understanding and the facts. *Philosophical Studies*, 132(1), 33–42.
- Elgin, C. (2017). *True enough*. Cambridge: MIT Press.
- Frigg, R., & Nguyen, J. (forthcoming). Mirrors without warnings. *Synthese*, 1–21.
- Greco, J. (2014). Episteme: Knowledge and understanding. In Timpe, & Boyd (Eds.), *Virtues and their vices*. Oxford: Oxford University Press.
- Grimm, S. (2006). Is understanding a species of knowledge? *The British Journal for the Philosophy of Science*, 57(3), 515–535.
- Hills, A. (2016). Understanding why. *Noûs*, 50(4), 661–688.
- Kelp, C. (2015). Understanding phenomena. *Synthese*, 192, 3799–3816.
- Kelp, C. (2017). Towards a knowledge-based account of understanding. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding* (pp. 251–271). New York: Routledge.
- Khalifa, K. (2013). Is understanding explanatory or objectual? *Synthese*, 190(6), 1153–1171.
- Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*. Cambridge: Cambridge University Press.
- Kvanvig, J. L. (2003). *The value of knowledge and the pursuit of understanding*. Cambridge: Cambridge University Press.
- Lawler, I. (forthcoming). Scientific understanding and felicitous legitimate falsehoods. *Synthese*, 1–29. <https://doi.org/10.1007/s11229-019-02495-0>
- Le Bihan, S. (forthcoming). Partial truth versus felicitous falsehoods. *Synthese*, 1–22. <https://doi.org/10.1007/s11229-019-02413-4>.

- Lipton, P. (2009). Understanding without explanation. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 43–63). Pittsburgh: University of Pittsburgh Press.
- Mizrahi, M. (2012). Idealizations and scientific understanding. *Philosophical Studies*, 160(2), 237–252.
- Mizrahi, M. (2017). Scientific progress: Why getting closer to truth is not enough. *International Studies in the Philosophy of Science*, 31(4), 415–419.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge: MIT Press.
- Nawar, T. (forthcoming). Veritism refuted? Understanding, idealizations, and the facts. *Synthese*, 1–19. <https://doi.org/10.1007/s11229-019-02342-2>.
- Niiniluoto, I. (2014). Scientific progress as increasing verisimilitude. *Studies in History and Philosophy of Science Part A*, 46, 73–77.
- Parker, G. (1978). Searching for mates. In J. Krebs & N. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 214–244). Oxford: Blackwell.
- Parker, G. (2017). Scientific progress: Why getting closer to truth is not enough. *International Studies in the Philosophy of Science*, 31(4), 415–419.
- Popper, K. R. (1963). *Conjectures and refutations*. London: Routledge.
- Popper, K. R. (1976). A note on verisimilitude. *The British Journal for the Philosophy of Science*, 27(2), 147–159.
- Potochnik, A. (2017). *Idealization and the aims of science*. Chicago: University of Chicago Press.
- Pritchard, D. (2009). Knowledge, understanding and epistemic value. *Royal Institute of Philosophy Supplement*, 64, 19–43.
- Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.
- Rancourt, B. T. (2017). Better understanding through falsehood. *Pacific Philosophical Quarterly*, 98(3), 66.
- Rice, C. (2015). Moving beyond causes: Optimality models and scientific explanation. *Noûs*, 49(3), 589–615.
- Rice, C. (2019). Models don't decompose that way: A holistic view of idealized models. *British Journal for the Philosophy of Science*, 70(1), 179–208.
- Rice, C. (forthcoming). Understanding realism. *Synthese*, 1–25. <https://doi.org/10.1007/s11229-019-02331-5>.
- Ross, L. D. (2020). Is understanding reducible? *Inquiry: An Interdisciplinary Journal of Philosophy*, 63(2), 117–135.
- Ross, L. D. (forthcoming). How intellectual communities progress. *Episteme*. <https://doi.org/10.1017/epi.2020.2>.
- Rowbottom, D. P. (2015). Scientific progress without increasing verisimilitude: In response to Niiniluoto. *Studies in History and Philosophy of Science Part A*, 51, 100–104.
- Saatsi, J. (2019). What is theoretical progress of science? *Synthese*, 196(2), 611–631.
- Salmon, W. C. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science*, 64, 461–477.
- Sliwa, P. (2015). IV—Understanding and Knowing. *Proceedings of the Aristotelian Society*, 115(1pt1), 57–74.
- Sliwa, P. (2017). Moral understanding as knowing right from wrong. *Ethics*, 127(3), 521–552.
- Sober, E. (2000). *The philosophy of biology* (2nd ed.). Boulder: Westview.
- Srinivasan, A. (2020). Radical externalism. *Philosophical Review*, 129(3), 395–431.
- Strevens, M. (2008). *Depth: An account of scientific explanation*. Harvard: Harvard University Press.
- Strevens, M. (2012). Précis of depth. *Philosophy and Phenomenological Research*, 84(2), 447–460.
- Strevens, M. (2013). No understanding without explanation. *Studies in History and Philosophy of Science*, 44, 510–515.
- Sullivan, E., & Khalifa, K. (2019). Idealizations and understanding: Much ado about nothing? *Australasian Journal of Philosophy*, 97(4), 673–689.
- Warenski, L. (forthcoming). Epistemic norms: Truth conducive enough. *Synthese*, 1–21. <https://doi.org/10.1007/s11229-019-02242-5>.
- Zach, R. (1978). Selection and dropping of whelks by Northwestern Crows. *Behaviour*, 67(2), 134–148.
- Zach, R. (1979). Shell dropping: Decision-making and optimal foraging in Northwestern crows. *Behaviour*, 68(1/2), 106–117.
- Zagzebski, L. (2001). Recovering Understanding. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue* (pp. 235–251). New York: Oxford University Press.

---

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.